

Text Recognition	Compare Packages			
	Feature / Aspect	PaddleOCR	Tesseract OCR	CnOCR
	License	Apache 2.0	Apache 2.0	Apache 2.0 (without Ultralytics)
	Year Released	2020 (v1), 2023 (v3)	2006 (v1), 2018 (v4 with LSTM)	2019
	Architecture	Deep learning-based (CNN + RNN with attention mechanisms)	Traditional OCR with LSTM-based recognition	Deep learning-based (CRNN architecture)
	Language Support	Supports multiple languages, with particularly strong performance on Asian scripts	Supports over 100 languages, making it highly versatile	Supports Simplified and Traditional Chinese (partial), English, and vertical text recognition
	Customization	Offers modular components for detection and recognition; supports fine-tuning	Supports custom language model training; offers extensive configuration options	Provides simple commands for training custom models
	Community Support	Rapidly growing community with active development	Large, established community with extensive documentation	Active development with a focus on Chinese OCR applications
	Accuracy	High accuracy, especially on complex layouts and Asian text	High accuracy on clean, printed text; may struggle with complex layouts	High accuracy in Chinese OCR; includes over 20 pre-trained models for various use cases
	Speed	Fast processing, optimized for real-time applications	Moderate speed; performance can vary based on preprocessing and image quality	Fast inference, suitable for real-time applications
	Best For	Asian text recognition	- Legacy document digitization - Multilingual projects (Latin/European scripts)	Chinese
	Deployment Support	Supports export to PyTorch and ONNX	Currently does not support export to other formats	ONNX format is the default export format
	TL;DR			
	If you need ONNX deployment, avoid Tesseract; otherwise:			
	If working on Latin or European scripts - TesseractOCR			
	If working on Asian languages other than Chinese - PaddleOCR			
	If working on Chinese - CnOCR			